

REPORT DOCUMENTATION PAGE			Form Approved OMB No. 0704-0188	
Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing this collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden to Department of Defense, Washington Headquarters Services, Directorate for Information Operations and Reports (0704-0188), 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to any penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number. PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.				
1. REPORT DATE (DD-MM-YYYY) 14/06/2012		2. REPORT TYPE Final		3. DATES COVERED (From - To) 15/08/2011 to 14/05/2012
4. TITLE AND SUBTITLE (STTR PHASE 1) Tools for Modeling & Simulation of Molecular and Nanoelectronics Devices			5a. CONTRACT NUMBER FA9550-11-C-0086	
			5b. GRANT NUMBER	
			5c. PROGRAM ELEMENT NUMBER	
6. AUTHOR(S) Mr. Anthony Waitz Prof. Jerzy Bernholc Prof. Kurt Stokbro			5d. PROJECT NUMBER	
			5e. TASK NUMBER	
			5f. WORK UNIT NUMBER	
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Atherton Quantum Insight, LLC 22 Holbrook Ln. Atherton, CA 94027 QuantumWise A/S, Lersø Parkallé 107, DK-2100			8. PERFORMING ORGANIZATION REPORT NUMBER 0001AC	
9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES) AF Office of Scientific Research 875 N. Randolph St. Room 3112 Arlington VA 22203			10. SPONSOR/MONITOR'S ACRONYM(S) AFOSR	
			11. SPONSOR/MONITOR'S REPORT NUMBER(S) AFRL-OSR-VA-TR-2012-1182	
12. DISTRIBUTION / AVAILABILITY STATEMENT •Distribution A: Approved for public release; distribution unlimited.				
13. SUPPLEMENTARY NOTES				
14. ABSTRACT Report developed under STTR contract for topic AF10-BT21. We have investigated the development of new multiscale methods that combine atomic-scale modeling of the active part of nano-sized electronic devices with continuum models of the surrounding passive parts. The methods are based on highly parallel algorithms to allow for simulations of realistic device structures. In Phase II, these methods will be developed and will be integrated into a new multiscale simulation module to the ATK package from the company QuantumWise. The resulting simulation package will be validated by a consortium of leading nanoscale device simulation groups in academia and industry, and used to address important problems in nano-science and semiconductor research and development. Additionally to investigating the feasibility of the Phase II activities, in Phase I we implemented a mechanism to allow the integration of other simulation engines into the ATK package. This allows these codes, which are all of academic origin to date, to be driven from within the ATK GUI, thereby making them far easier to access.				
15. SUBJECT TERMS STTR Report				
16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT None	18. NUMBER OF PAGES 13
a. REPORT Not classified	b. ABSTRACT Not classified	c. THIS PAGE Not classified		
				19b. TELEPHONE NUMBER (include area code) (650) 269-4566

Final Technical Report: Tools for Modeling & Simulation of Molecular and Nanoelectronics Devices

Project Objectives

The goal of this STTR funded project is to overcome some of the significant obstacles to modeling the electrical properties of nano-scale devices by implementing new high-performance multiscale modeling methods. The team consists of Atherton Quantum Insight LLC (PI), North Carolina State University, and QuantumWise A/S. In Phase I of the STTR, the focus is on comparing existing codes already developed by NCSU and QuantumWise, investigating technical approaches to various problems, creating a plan for Phase II, and marketing outreach for the new technology.

Work Performed

Phase I was broken into eight deliverables as specified in the original proposal. These are shown in Table 1 below with the addition on one new deliverable (6a). In "Status Report 1", delivered October 2011, we described the completion of the first three deliverables which were all related to selecting and preparing a finite element (FE) framework approach which would be used in implementing the multiscale capability in Phase II. In status Report 2 we described the completion of the next four deliverables (nos. 4, 5, 6, and 6a) in Table 1 below. In this final report we discuss deliverable no. 8 and further activities related to the FE work already done.

No.	Deliverable	Status Report 1	Status Report 2	Final Report
1	Selection of FE libraries to use.	Delivered		
2	Implementation of generation routing for adaptive FE grid as obtained in the NanoPar project.	Delivered		
3	Common data file format for visualizing FE grids.	Delivered		
4	Review and analysis of the algorithms and methodologies used in the NCSU and QuantumWise ATK NEGF transport codes.		Delivered	
5	Verification by explicit comparison of results generated by at least two completely independent codes for a test suite of explicit device configurations.		Delivered	
6	Initial geometry-passing interface between the academic and ATK codes implemented using a software "plug-in" mechanism.		Delivered	
6a	Plans for "plug-in" type interface to facilitate output passing between the academic and ATK codes, which will enable easy analysis of the results generated by the academic from within ATK.		Delivered	
7	Development of a detailed plan of methodology and algorithm integration.			Delivered
8	Detailed plan for marketing of the future capabilities, identification of current and future customers, buildup of customer relations.			Delivered

Table 1 - Deliverables

Results Obtained

Deliverable No. 1 – Selection of FE Libraries to Use

We have implemented a prototype DFT simulation software using two different open source Finite Element (FE) libraries: DEALII and FENICS. These two libraries have been compared in terms of functionality and performance. The study clearly shows that the DEALII library has the best performance and will best fit our purpose. See “Appendix A - Finite Element Libraries Comparison” for the details. There still are a number of issues with the library which need to be addressed before it can be used in commercial software and these issues has been forwarded to the developers of the library.

Deliverable No. 2 – Implementation of Generation Routing for Adaptive FE Grid as Obtained in the NanoPar Project

In order to be able to evaluate the FE libraries we have implemented a routine for generating FE grids for atomic-scale geometries. The generation of the FE grid is based on division of the space, such that each grid element contains the same amount of electron density.

Deliverable No. 3 – Common Data File Format for Visualizing FE Grids

As part of the implementation of the prototype FE software we have implemented a data structure for the FE grids. The data structure is based on the internal data structure of the FE libraries. Our white paper study shows visualizations of the FE grids stored using the data structure.

Deliverable No. 4 - Review and analysis of the algorithms and methodologies used in the NCSU and QuantumWise ATK NEGF transport codes

Both ATK and NCSU codes use the same formula of non-equilibrium Green functions (NEGF) as described in the literature [1]. The most time-consuming part in the NEGF formalism is to calculate the charge density matrix,

$$D_{\mu\nu} = \int_{-\infty}^{\infty} d\varepsilon [\rho_{\mu\nu}^L(\varepsilon)n_F(\varepsilon - \mu_L) + \rho_{\mu\nu}^R(\varepsilon)n_F(\varepsilon - \mu_R)] , \quad (1)$$

where $n_F(\varepsilon - \mu_{L,R})$ are the Fermi-Dirac distribution functions and $\rho_{\mu\nu}^{L,R}(\varepsilon) = \frac{1}{\pi} [G(\varepsilon)\Sigma_{L,R}(\varepsilon)G^\dagger(\varepsilon)]_{\mu\nu}$ are the left/right spectral density matrices. The self-energy operators $\Sigma_{L,R}$ are used to describe the semi-infinite left and right leads and the Green function G is calculated by

$$G(\varepsilon) = [\varepsilon S - H - \Sigma_L - \Sigma_R]^{-1} . \quad (2)$$

The charge density is calculated by

$$\rho(r) = \sum_{\mu\nu} \phi_\mu(r) D_{\mu\nu} \phi_\nu(r) , \quad (3)$$

where $\phi_{\mu,\nu}(r)$ denote the basis set, which must consist of localized orbitals to result in finite-size expansion of the Green's function operator, eq. (2).

Although the same formulas are used in both the ATK and NCSU codes, the implementations are different in the choice of a basis set, parallelization procedures, input parameters, etc. In the following, we describe the differences between the two codes.

Basis Set

In the NCSU code, the localized orbitals are optimized variationally for the system under consideration [2]. Its accuracy is controlled by the radii and number of the localized orbitals. Since the orbitals are optimized for the specific system, one can obtain a small but nearly-optimal basis set for the required accuracy. For example, four to six orbitals per carbon atom are usually good enough in the transport calculations and for absolutely-converged total energy. The disadvantage of this basis set is the need to optimize the orbitals for each system, which sometimes takes a substantial number of iterations, and the sizable radii of the orbitals.

In ATK code, the localized orbitals are the solutions of spherical symmetric confinement potential. Different parameters of the confinement potential can be varied to obtain an optimal basis set for some reference system. The ATK comes with a number of generic basis sets for each element [3]. The basis sets are divided into Single Zeta, Double Zeta and polarization orbitals. The Zeta orbitals are the valence orbitals for angular momenta shells which are occupied for the atom, while the polarization orbitals are the first angular momentum shell that is unoccupied. This basis set can be significantly larger than the optimal one, but it is more easily transferable between the different systems.

Parallelization

The NCSU code implements multi-level parallelization using the message passing interface (MPI). When the charge density matrix, eq. (1), is calculated, the sampling of the energy points in the integration can be a few hundreds or even more in the case of large, non-equilibrium bias. For each energy point, one needs to invert a matrix to get the Green function in eq. (2) and to perform matrix multiplications to obtain the spectral density matrix. The N_{total} MPI processes are partitioned into subgroups with a two dimensional Cartesian topology $N_{\text{energy}} \times N_{\text{matrix}} = N_{\text{total}}$. Each subgroup with N_{matrix} processes performs the matrix operations (inversion and multiplication for a few energy points by calling a ScaLapack library). ScaLapack's data structure is used for distribution of all of the matrixes, including Green functions, self-energy, overlap and Hamiltonian matrixes. This is critical for a large scale calculations, since the memory required increases at least linearly with the size of the system. There is another level of parallelization for the current multi-core architecture, i.e., linking the multi-threaded ScaLapack library which is available on most of the supercomputers. For example, one can use 8 to 16 cores per MPI process on Cray XE6. This

parallelization strategy is suitable for employing thousands of processors, and test calculations have been carried out for over 3,000 atoms.

The ATK code is parallelized over k-points and energy points. The energy point parallelization is similar to the NCSU code, while the NCSU code currently only support Gamma point simulations at present (a single k-point), although a k-point implementation is in progress. The matrix operations for each energy point is performed using the MKL library in the ATK code through an OpenMP parallelization. Thus, similar to the NCSU code, the ATK code divides the processing units N_{total} into subgroups with a two dimensional Cartesian topology $N_{\text{energy}} \times N_{\text{matrix}} = N_{\text{total}}$. Mpi parallelization is performed over the N_{energy} processors while OpenMP parallelization is performed for the N_{matrix} processors.

After the charge density is self-consistently determined, the transmission coefficients are calculated by Landauer formula

$$T(\epsilon) = \text{Tr}(G\Gamma_L G^\dagger \Gamma_R). \quad (4)$$

Similarly to the charge density calculation, the transmission also needs the sampling of energy points and matrix operations at each energy point. We use the same parallelization scheme as in the previous discussion.

Deliverable No. 5 - Verification by explicit comparison of generated results

As initial verification, we present detailed comparisons between the results generated by the ATK and NCSU codes. These codes have been written completely independently and do not share any components. Furthermore, the results have been obtained on two different platforms: (i) an 8-core Linux server for the ATK code, and (ii) the Cray XT supercomputer for the NCSU code. To date, we have compared two systems: (i) a Stone-Wales defect in a graphene nanoribbon, and (ii) a molecular junction consisting of a benzene ring connected to gold leads via thiol linkages. While the first system involved only carbon atoms, the second one includes C, H, S, and Au atoms and a more complicated atomic structure. As will be shown in detail below, the results generated by the two codes compare well and produce the same findings.

Graphene Nanoribbon

We consider a graphene nanoribbon with a 5775 "Stone-Wales" defect as shown in Fig.1a. The left and right electrodes are ideal zigzag- edge nanoribbons, which have edge states near the Fermi energy. In the central scattering region, a 5775 defect (see the blue area in Fig. 1(a)) is introduced. The transmissions calculated by the NCSU and ATK codes are shown in Fig.1(b). The agreement is excellent, especially since two very different codes have been used. The sharp peak around the Fermi energy is due to an edge state, which is not dramatically affected by the defect. At other energy points, the transmission is reduced due to scattering from the defect. The small discrepancies between two curves may be due to the different basis sets and/or pseudopotentials used in the two calculations.

In the NCSU calculations with four, six, or nine orbitals per atom, the results are very similar with negligible differences. The effect of orbital radius is also very small, which was verified by comparing results for 3.7 Å and 4.2 Å radii.

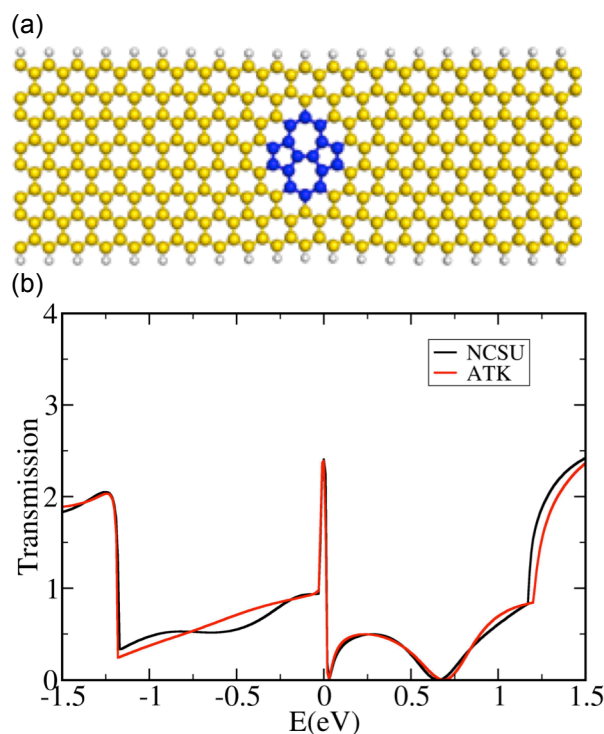


Fig.1. (a) Atomic structure of a zigzag graphene nanoribbon with a 5775 defect (blue area). (b) Transmission curve calculated by NCSU (black curve) and ATK (red curve) codes .

Molecular Junction

The atomic structure of a benzene ring attached by thiol linkers to gold nanowires is shown in Fig.2(a). The transmission coefficients calculated by both codes are shown in Fig. 2(b). In the NCSU-code calculations, three basis sets were explored. The first basis set uses 10 orbitals per Au atom and 6 orbitals per S, C, and H atoms with radii of 6.5 a.u. The second one uses same number of orbitals with larger radii of 8.5 a.u., and the third one includes 20 orbitals per Au atom and 9 orbitals per S, C, and H atoms with radii of 10.0 a.u. We find that the results obtained with the second basis set are converged and essentially coincide with those of the third basis set. The transmission calculated with the first basis set is different from the converged one, most significantly near the peaks at -1 eV and +2 eV, although the agreement around Fermi energy is reasonable. The ATK results with DoubleZetaPolarized basis set are different from the NCSU ones, while results obtained with the DoubleZetaDoublePolarized basis set are comparable to ones from the NCSU code, especially around the Fermi energy.

In order to obtain the I-V curve, one needs to calculate the transmission at different biases. In the following, we use the second basis set for NCSU and the DoubleZetaDoublePolarized basis set for ATK calculations. Fig. 3(a) shows the transmission at biases of 0.0 and 0.4 V. Apart from small discrepancies around the peaks at -1.0 and 2.0 eV, the agreement between the two codes is very good. The discrepancies might be mostly due to different pseudopotentials used by the different codes. The I-V curves are shown in Fig. 3(b) and the results from the two codes almost coincide, because the current is mainly determined by the transmission around the Fermi level.

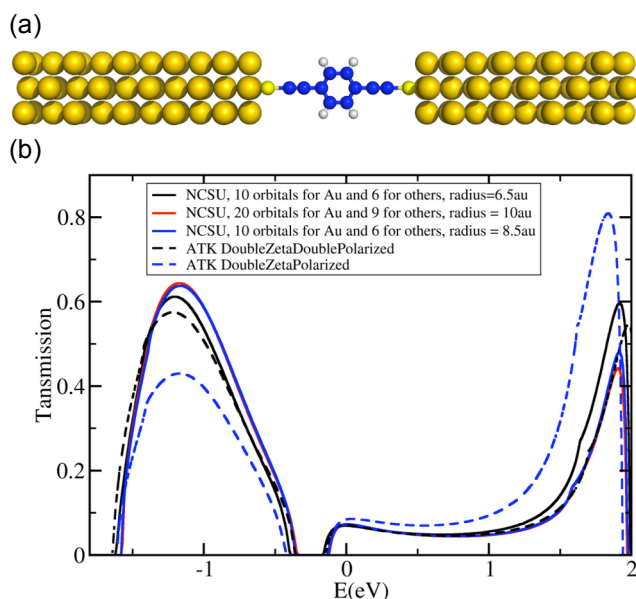


Fig.2. (a) Atomic structure of a molecule sandwiched by gold wires. (b) Transmission coefficient calculated by NCSU (solid lines) and ATK (dashed lines) codes with different basis sets.

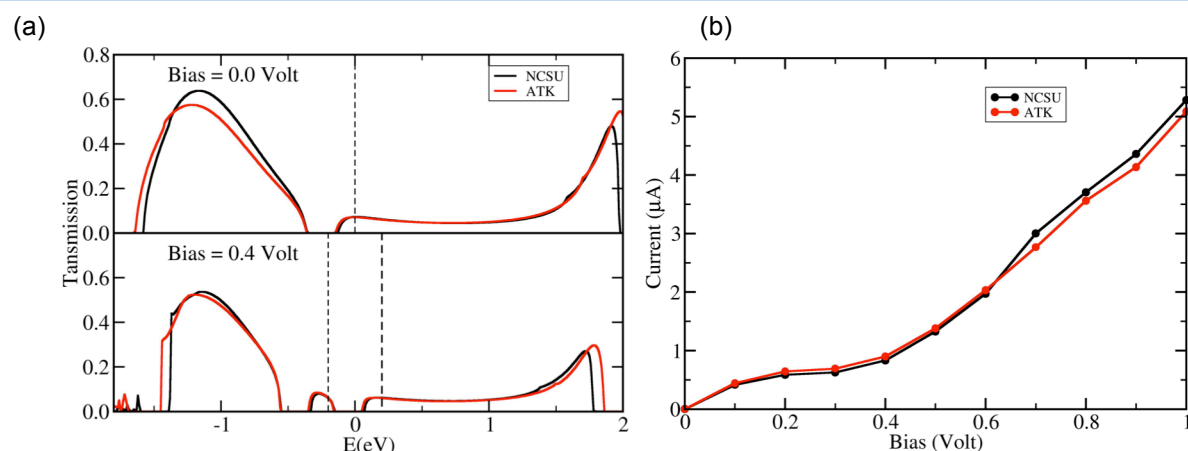


Fig.3. (a) Transmission at biases of 0.0 and 0.4 V from NCSU (black) and ATK (red) codes. (b) I-V curve for the system shown in Fig. 2a.

Deliverable No. 6 - Initial geometry-passing interface between the academic and ATK codes implemented using a software "plug-in" mechanism

We have developed a graphical user interface for the NCSU transport code based on the QW-developed external code plug-in. This interface makes it much easier to build a complex device structure employing the easy-to-use graphical tools of ATK. As a side-benefit of this effort, we realize that this plug-in mechanism can be used with other academic codes and can also be extended to handle outputs of such codes thereby giving access to many otherwise underutilized codes.

A graphical device configuration in ATK includes atomic structures and lattice parameters for the leads and the central scattering part. Once the device configuration is set up, all input parameters required by NCSU transport calculations can be controlled and modified by this interface. The interface includes several panels which control the initial setup, self-consistent (SCF) steps and accuracy, configuration of real space grids, species, etc. Fig. 1 shows two screenshots from the interface. From the panel "Grids" one can set the real space grids and processor topologies for DFT and NEGF calculations. From the panel "Species" one can choose pseudopotential files, the number of orbitals and the radii for each species. As the output, this interface creates all input files for NCSU calculations and the job scripts for a supercomputer (in our case usually the Cray XT), which is chosen in the panel "Setup".

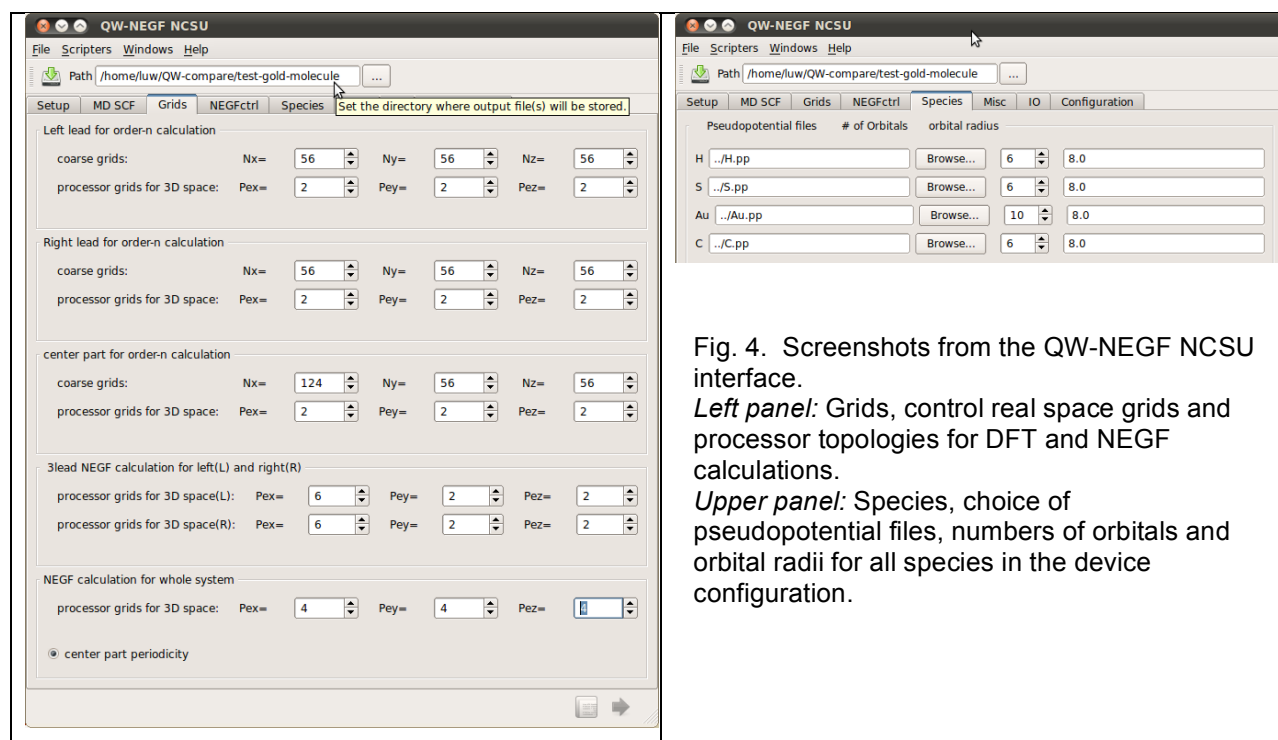


Fig. 4. Screenshots from the QW-NEGF NCSU interface.

Left panel: Grids, control real space grids and processor topologies for DFT and NEGF calculations.

Upper panel: Species, choice of pseudopotential files, numbers of orbitals and orbital radii for all species in the device configuration.

As a non-trivial example, we show below a nanotube-DNA-nanotube configuration that was assembled using an open-source DNA builder followed by device and input file setup in the ATK code. The completed configuration, consisting of 570 atoms, is being investigated on a Cray XE6 supercomputer.

The graphical user interface has dramatically enhanced the productivity of a mid-level graduate student at NCSU, who can now assemble the input files needed for complex DNA conductance simulations at a fraction of the time that he needed previously. The NCSU code requires several input files specifying the atomic coordinates of the leads and the central region, as well processor grid configurations for each. These can now be generated simultaneously, creating the complex combined geometries in one step. Furthermore, instead of calculating the initial atomic positions of the 3D structures by hand, they can now be obtained from a point-and-click visual interface.

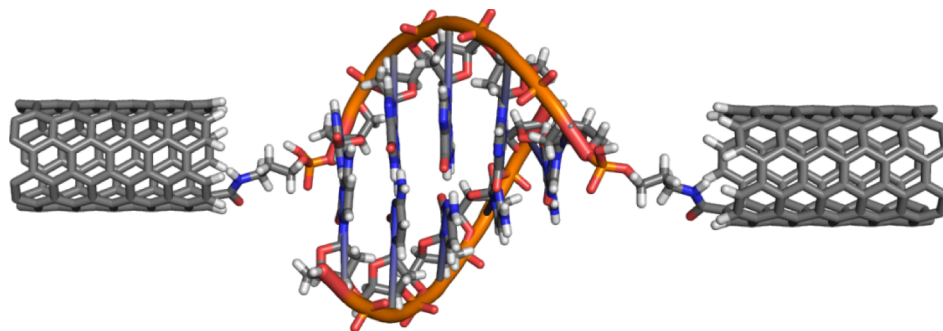


Fig. 5. Test configuration for conductance studies of matched and mismatched DNA pairs connected to nanotube leads. The calculations include 570 atoms.

In Phase II, we plan to generalize this plug-in structure to enable an easy to use interface for three-terminal configurations, to facilitate quick setup of realistic devices for investigation of current amplification ratios, leakage currents, and onset voltages.

Deliverable No. 6a - Plans for "plug-in" type interface to facilitate output passing between the academic and ATK codes

In a Phase II project QuantumWise will extend their platform such that ATK can recognize output data from external codes. The new addition will be through a plug-in mechanism, such that it is possible for third parties to develop the connections independently of QuantumWise. There will be two different approaches available for making the connection:

- Plug-ins that convert the data to a data structure recognizable by ATK, and ATK's analysis modules can then be used for performing the analysis.
- Plug-ins that recognize and operate directly on the third party generated data.

Another extension planned for the Phase II project will be a plug-in facility within the ATK job manager, which makes it possible to send and retrieve data between a laptop client and a supercomputer for an external code.

The extension will make it possible for NCSU to fully integrate their code into the ATK platform, i.e. setup the system, prepare input files, send the job of to a supercomputer and perform the analysis.

Deliverable No. 7 - Development of a detailed plan of methodology and algorithm integration

In order to simulate realistic nano-scale devices there is a need for models which can solve the following challenges:

1. Simulate systems with more than 10000 atoms in the active device region
2. Develop highly transferable, efficient yet accurate basis sets for very large scale calculations.
3. Improve the description of exchange and correlation in order to reproduce semiconductor band gaps.
4. Simulate systems with three or more current carrying electrodes,
5. Include the electrostatic effects of the surrounding control system which can be large compared to the active device region.

It is these challenges that will be addressed in the phase II of the project.

Since the quantum transport problem occurs in a linear geometry, it can be formulated as an $O(N)$ approach [5] and thus scale linearly with system size. Indeed, the NCSU group has already carried out calculations with over 3,000 atoms on a Cray XT4. With increased parallelization and interconnect speed, 10,000 atoms are imminently feasible. For example, the NCSU group already rewrote their standard real

space code (with delocalized basis) for large processor counts using pthreads and OpenMP, in addition to MPI. The rewritten code, where several bottlenecks have been eliminated, scales nearly linearly from 4k cores (150.2 secs) to 128k cores (6.9 secs) on the Cray XE6. However, the quantum transport calculations must use localized orbitals in order to localize the expansion of the Green's function. The $O(N)$ code has not yet been enhanced, but we expect similarly improved parallelization and linear scaling to well over 10,000 atoms. Several of the routines and procedures are shared between the two NCSU code bases, making this part of the project a low-risk endeavor. The applicable NCSU parallelization strategies will be transferred to ATK for incorporation into its code base.

The development of a transferable, optimal basis set that is easy to deploy is a research question that is outside of the scope of Phase I STTR, but it will be addressed in Phase II. Specifically, since the NCSU approach generates system-specific optimized orbitals, we will investigate whether one can generate "coordination-optimized" orbitals, e.g., for 2-, 3-, and 4-coordinated carbon atoms, which would lead to higher accuracy and thus fewer orbitals for convergence, while still being transferable to different systems. The NCSU approach also generates matrix elements in the essentially minimal basis of optimal orbitals. For regions far from the potential drop, these elements should not change and one would defacto have a high accuracy tight-binding basis, far better than in existing "ab initio" tight-binding approach. The resulting "hybrid" tight-binding DFT approach should be able to handle very large systems while maintaining full DFT accuracy.

Appropriate description of semiconductor band gaps is important in quantitative simulation of devices. In Phase II of the project we will incorporate the modified Becke-Jones functional [6], which reproduces well the band gaps of advanced semiconductors. We will also explore the incorporation of screened exact exchange, which should not greatly increase the computation time if proper localization strategies are employed.

The multi-terminal self-consistent formulation was previously developed by Bernholc, Lu and collaborators [7-9]. This methodology will be made more accessible and implemented in ATK in Phase II, enabling routine studies of transistor structures by government and industrial researchers. New capabilities will also concern simple circuits, which could be fully implemented if sufficient computer power is available.

Description of electrostatic effects outside of the active device structure is key to faithful modeling of emerging quantum-based electronic and sensing circuitry. We will implement a finite element description, which allows for a multi-scale model, where the active device region is described with a fine mesh and the surrounding control system is described with a coarser mesh. Such an implementation will require a good finite element library which can be integrated with ATK.

In the first part of this Phase I project we investigated two different candidate finite element libraries, DEAL II and FENICS. Although both libraries were promising, our investigation showed that each of the libraries were missing a few features needed for our purpose [10]. The report has now been sent back to the authors of these libraries and we are in dialogue with the groups regarding our requirement for further development. Since we are able to address these problems at a very early stage we are in a good position for a successful implementation of a multi-grid model in a Phase II project.

An alternative approach is to use parallel multigrid techniques. NCSU has a highly parallel Poisson multigrid solver that is routinely used in hybrid DFT/orbital-free-DFT simulations of solvated biomolecules. These systems routinely consist of over 10,000 atoms, yet this solver consumes only a very small fraction of the overall computer time. If the finite-element approach runs into technical difficulties, we will experiment with the multigrid solver, which could be modified to handle different grid densities in the passive and active device regions. As a technical note, a multigrid Poisson equation solver is stable when different grid resolutions are used in different parts of the solution domain. A multigrid-based eigenvalue solver is not and special techniques have to be used to ensure stability, reducing effectiveness of the iterations.

Deliverable No. 8 - Detailed plan for marketing of the future capabilities, identification of current and future customers, buildup of customer relations

Commercial Potential and Market Requirements if Project is Carried Through Phase II

According to a GP Bullhound report [4], the market size for modeling software for nanoscale electronics was \$110 M in 2009 and was growing at a rate of 32%. Perhaps these estimates are overly optimistic, but we know, from personal experience, that the market for such software is expanding rapidly and presents a good opportunity for adoption of products that satisfy the ever escalating needs of customers in this domain.

The largest company in the quantum-accurate simulation market is Accelrys. Their focus is primarily on drug discovery with a secondary focus on general materials science. We know, from talking with Accelrys employees, that Accelrys considered entering the quantum-accurate nanoscale electronics simulation market but decided not to because the focus was too different from their current customer base.

The traditional chip design software companies, Synopsys, Cadence, and Mentor Graphics, in the industry known as Electronic Design Automation (EDA), are approaching the quantum-accurate nanoscale electronics simulation market from the top-down. Currently, none of those companies have quantum-accurate simulation products – but eventually they must.

With Accelrys and the EDA companies not participating in the quantum-accurate nanoscale electronics simulation market, a window of opportunity has been formed for new entrants that can satisfy the needs of this growing opportunity. We believe that if we are able to create what the “power user” requires (see definition of Power User Requirements in box below), then we will be in a position to command a significant portion of the emerging market for nanoelectronic simulations. From a high level perspective, this is what we would propose if invited to apply for Phase II of the STTR.

A nanoelectronics simulation platform that has: (i) the ability to seamlessly couple atomistic and mesoscopic regimes, (ii) a multi-terminal capability to enable full evaluation of device structures, (iii) a highly parallelized architecture to leverage modern supercomputers, (iv) the ability to model many-body effects beyond standard DFT approaches, and (v) been validated by explicit comparison to experimental data.

Power User Requirements for a Defense-oriented Platform for Simulation of Nanoelectronics

Because of the magnitude of this opportunity, QuantumWise A/S plans to establish a US subsidiary if the Phase II portion of this project is funded. QuantumWise would essentially be co-investing alongside AFOSR in the Phase II portion of this endeavor, which would ensure rapid success of this ambitious project and its subsequent adoption by government labs, industry and academia.

Identification of Potential Customers and Development of Relationships

During Phase I of this STTR we were able to reach out to many potential customers and communicate our intent to build a nanoelectronics simulation platform that meets the Power User Requirements described above. We did this by attending conferences, visiting customers individually, and participating in a pre-competitive academic research consortium. For all organizations doing any work that involved nano electronics, the response has always been positive – most organizations either had an immediate need or anticipated a need in the near future.

We attended the 2011 NanoTechnology for Defense Conference in Seattle. There, in one-to-one meetings with various prime defense contractors, the idea of a defense-oriented platform for simulation of nanoelectronics was presented. In general the idea was very well received, even though at that time it was in a nascent form without a demonstration version available – not even the integration of the NCSU code into ATK. The companies that we had one-to-one meetings with were: BAE, Boeing, Goodyear, Lockheed Martin, Rolls Royce, and Northrop Grumman. Outside of this conference we also met with the defense company Raytheon. Although the needs of these defense companies are diverse, one general

theme is the ability to simulate actual whole devices or structures, another is throughput of simulation. Credibility of the results were also important as would be expected. This helped form our Power User Requirements.

We also met with a number of federal labs including: AFLR, NRL, ARL, NIST, Argonne, and NREL. Similar to the defense contractors, we found that these organizations were interested in a wide variety of technical problems, but in general, their needs aligned with our Power User Requirements shown above.

We also attended the IEDM and SYSPAD conferences which are both focused on semiconductors. In both conferences we met with many physicists and materials scientists from semiconductor companies. With these potential customers, the requirements are more narrowly defined to the realm of semiconductor devices. In general there is a feeling within this community that the need for quantum-accurate simulations is becoming a reality and that relying on their existing, non-quantum based solution, is becoming serious bottleneck to progress. Also this group emphasized the need for high throughput tools. This industry is not accustomed to waiting hours to simulate a semiconductor structure. For this market, a high degree of parallelism will be very valuable.

The Market Need for a Nanoelectronics Simulation Platform

Additionally, our out-reach efforts have included approaching a handful of leading academics groups, besides Prof Bernholc's group, to gauge interest in including their codes within the ATK platform. We have universally found a high level of interest and are in talks with a number of these groups on the specifics of such collaborations.

In Phase II of this project, one of the activities would be to focus on finding codes that were developed within the DoD labs that could benefit from integration into the ATK platform using the newly developed plug-in model. We believe that this would be of great use to the labs in allowing a wider user base to access the codes that have already been developed – thereby leveraging the existing DoD investment. This work will be done in parallel with enhancing and extending the technical capabilities of the existing codes in the manner that was described above.

Plan for Marketing of Future Capabilities Created by Phase II

We believe that in Phase I of this project, we came to a good understanding of the needs of the defense and related communities in the area of quantum-accurate nanoscale electronics simulation. Such simulations appear necessary to maintain Moore's law beyond the limits of silicon technology. If implemented at projected speeds, it would dramatically enhance the speed of data processing in the battlefield, improve smart weapons, sensors, and information processing in general. Nanoscale devices are being conceived as the path to future electronics with ultra-dense, ultra-fast molecular-sized components, with very small power requirements and persistent, reprogrammable memories. Self-assembly of nanoscale devices, potentially using bio-inspired nanoscale processes, is also being envisioned as an avenue for overcoming the second Moore's law, namely that the cost of electronic devices is inversely proportional to their density. Such devices would be exceptionally fast and non-volatile, while consuming very little power. They would find many uses in intelligent projectiles, countermeasures, and autonomous systems, as well as dramatically enhance the speed of data processing in the battlefield, improve smart weapons, sensors, and information processing in general. Various nanoscale-based experimental logic and memory elements have been fabricated already, although not yet with methods suitable for mass production. to address the kinds of problems that are being faced.

If we are awarded a Phase II project, we would continue the outreach and marketing activities with the intent of getting early adopters for our newly developed technology. In addition to the marketing and sales activities that one would expect for promoting new software, we would also hold day-long seminars at the DOD Labs and other relevant Federal Labs. The goal of these seminars would be to: deliver academic training in the theory and need for quantum transport calculations, illustrate the types of problems that can be solved with such technology, and deliver a mini-training on the software to give prospective users the chance to get the feel of using it.

Another part of the marketing plan for the new technology, we would be present at relevant conferences, such as the Defense for Nanotechnology, and the NSTI Nanotech conferences. We would also give papers at these conferences that would be based on the new technology developed in Phase II. Additionally we would have a booth/table at some of the conferences where anticipate the turnout to be significant.

Estimates of Technical Feasibility & Future Plans

In the project we have investigated the feasibility of the following technologies.

Parallel algorithms for transport

The main bottleneck in the transport calculations is the calculation of the non-equilibrium Green's functions. The Green's function is calculated by inversion of a block diagonal matrix at a number of energy and k-points. Both codes are parallelized over energy and k-points. The ATK code use a more efficient block diagonal solver than NCSU, while the NCSU code solve the block diagonal inversion in parallel giving a higher parallelism. In a Phase II we will integrate the two approaches to obtain a highly efficient and parallel scalable algorithm.

Basis sets for describing the electronic structure

We have compared the accuracy of the NCSU and the ATK code. The comparison shows that the NCSU code can obtain the same accuracy as the ATK code with fewer orbitals. In a Phase II we will investigate in more detail the difference between the NCSU and ATK basis sets, and develop special ATK basis sets for important systems.

Finite element grids

We have tested different finite element grids and made a prototype implementation of a finite element code. In a Phase II the prototype will be rolled out into a released version.

Multi terminal devices

Prior to the Phase I, we made a prototype of a multi-terminal transport code. This multi-terminal capability will be rolled out into ATK in Phase II.

Appendix A - People Involved

Mr. Anthony Waitz, Atherton Quantum Insight, LLC
Dr. Sergey Barabash, Atherton Quantum Insight, LLC
Prof. Jerzy Bernholc, North Carolina State University
Prof. Wenchang Lu, North Carolina State University
Prof. Miroslav Hodak, North Carolina State University
Mr. Bikan Tan, North Carolina State University
Prof. Kurt Stokbro, QuantumWise A/S
Mr. James Emil Avery, University of Copenhagen

Appendix B - References

- [1] M. Brandbyge, J. L. Mozos, P. Ordejón, J. Taylor, and K. Stokbro, "Density-functional method for nonequilibrium electron transport", *Physical Review B*, vol. 65, no. 16, 2002.
- [2] J. L. Fattebert and J. Bernholc, "Towards grid-based $O(N)$ density-functional theory methods: Optimized nonorthogonal orbitals and multigrid acceleration", *Physical Review B*, vol. 62, no. 3, pp. 1713-1722, 2000.

- [3] The ATK reference manual,
<http://www.quantumwise.com/documents/manuals/latest/ReferenceManual/index.html/>
- [4] GP Bullhound, Company Report, Atomistix A/S, Nanoscale EDA, October 2006. Unpublished investment firm report.
- [5] Nardelli MB, Fattiberto JL, Bernholc J. $O(N)$ real-space method for ab initio quantum transport calculations: Application to carbon nanotube-metal contacts. *Physical Review B* 2001; 64(24):245423.
- [6] Tran F; Blaha P. 2009. Accurate Band Gaps of Semiconductors and Insulators with a Semilocal Exchange-Correlation Potential. *Phys Rev Lett.* 102:226401.
- [7] Saha K, Lu W, Bernholc J, Meunier V. First-principles methodology for quantum transport in multiterminal junctions. *J. Chem. Phys.* 131, 164105 (2009).
- [8] Saha K, Lu W, Bernholc J, Meunier V. Electron transport in multiterminal molecular devices: A density functional theory study. *Phys. Rev. B* 81, 125420 (2010).
- [9] Saha, K.K., Nikolic, B.K., Meunier, V., Lu, W. & Bernholc, J. Quantum interference controlled three-terminal graphene-nanoribbon--ring-shaped-molecule transistors. *Phys. Rev. Lett.* 105, 236803 (2010).
- [10] Comparison between two finite element libraries, Deal.II and FEniCS/dolfin, for use as back-ends to a finite element DFT in ATK, Quantum Insight and QuantumWise A/S, October 2011.